

Recognition of Handwritten Digits using Proximal Support Vector Machine

A thesis submitted in partial fulfilment of the
requirements for the degree of

Master of Technology

in

Electronics and Communication Engineering
(Signal and Image Processing)

by

Swapna Prava Ekka

Roll No. 212EC6186



Department of Electronics and Communication Engineering

National Institute of Technology Rourkela-769008

2014

Recognition of Handwritten Digits using Proximal Support Vector Machine

A thesis submitted in partial fulfilment of the
requirements for the degree of
Master of Technology
in
Electronics and Communication Engineering
(Signal and Image Processing)

by

Swapna Prava Ekka

Roll No. 212EC6186

under the guidance of

Dr. Samit Ari



Department of Electronics and Communication Engineering

National Institute of Technology Rourkela-769008

2014

Declaration

I hereby declare that the work presented in the thesis entitled “*Recognition of Handwritten Digits using Proximal Support Vector Machine*” is a bonafide record of the research work done by me under the supervision of Prof. Samit Ari, Department of Electronics & Communication Engineering, National Institute of Technology, Rourkela, India and that no part thereof has been presented for the award of any other degree.

Swapna Prava Ekka

Roll No: 212EC6186

Dept. of Electronics & Comm. Engg.

National Institute of Technology

Rourkela, India-769 008

Certificate

This is to certify that the thesis entitled, “*Recognition of Handwritten Digits using Proximal Support Vector Machine*” submitted by Ms. Swapna Prava Ekka partial fulfillment of the requirements for the award of Master of Technology Degree in Electronics and Communication Engineering with specialization in “Signal and Image Processing” during session 2013-2014 at the National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance. To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University/ Institute for the award of any degree or diploma.

Dr. Samit Ari

Dept. of Electronics & Comm. Engg.

National Institute of Technology

Rourkela, India-769 008

Acknowledgements

I would like to make my deepest appreciation and gratitude to Prof. Samit Ari for his valuable guidance, and encouragement during the course of this project.

My special thank goes to Prof. S Meher Head of the Department of Electronics and Communication Engineering, NIT, Rourkela, for providing us with best facilities in the Department and his timely suggestions.

I want to thank all my teachers Prof. L.P. Roy, Prof A.K Sahoo, and Prof U.K .Sahoo for providing a good background for my studies and research thereafter.

I would also like to extend my thanks to all my classmates and lab mates for their time, invaluable suggestions and help. I thank my parents and family members for the support, encouragement and good wishes, without which I would not have been able to complete my thesis.

Swapna Prava Ekka

Roll No: 212EC6186

Abstract

Handwritten Digit Recognition System involves reception and interpretation of handwritten digits by a machine. Due to variation in shape and orientation of handwritten digits, it is difficult for a machine to interpret handwritten digits. Handwritten digit Recognition has a wide area of research due to its vast applications like automatic bank cheques processing, billing and automatic postal service. In this thesis, an Offline Handwritten Digit Recognition System is presented. The recognition system is broadly divided into 2 parts, first part is feature extraction from handwritten images and the second one is classification of feature vector into digits. We propose descriptors for handwritten digit recognition based on Histogram of Oriented Gradient (HOG) feature .It is one of the widely used feature vector for object detection in computer vision. For classification of features, linear Proximal Support Vector Machine (PSVM) Classifier is proposed. This is a binary class classifier which is further converted to a 10 class classifier by means of One against all algorithm. Due to small training time, PSVM classifier is preferable over standard Support Vector Machine (SVM) Classifier. The handwritten images both for training and testing are taken from MNIST database. The performance of the system is measured in terms of Sensitivity, Accuracy, Positive Predictivity and Specificity. The performance of PSVM classifier is better compared to Artificial Neural Network(ANN).

Keywords- Histogram of Oriented Gradient (HOG) feature, Proximal Support Vector Machine (PSVM) Classifier, one against all

Contents

Abstract	IV
List of Figures	VII
List of Tables	VIII
1. Introduction	1
1.1 Introduction	2
1.2 Literature Review	3
1.3 Motivation of Thesis	3
1.4 Brief introduction to Supervised Learning	4
1.5 System Overview for Handwritten Digit Recognition	6
1.6 Dataset Description	7
1.7 Thesis Outline	8
2. Histogram of Oriented Gradient (HOG) Based Feature Extraction	9
2.1 Introduction	10
2.2 Steps for Feature Extraction	11
2.2.1 Gradient Computation	11
2.2.2 Block Division	12
2.2.3 Orientation Binning	12
2.2.4 Block Description	13
2.2.5 Block Normalization	13
2.3 Feature Dimension Calculation	14
2.4 Experimental Results and Discussion	15
2.4.1 Block Size	15

2.4.2	Orientation Bin	16
2.4.3	Calculation of Feature Dimension	17
2.5	Conclusion	18
3.	Multiclass- Linear Proximal Support Vector Machine (PSVM)	
	Classification	19
3.1	Introduction	20
3.2	Linear Proximal Support Vector Machine for Binary Class	20
3.3	Implementation of Multiclass PSVM	23
3.3.1	One-against-all	24
3.4	Confusion Matrix	26
3.4.1	Confusion Matrix for Binary Class	26
3.4.2	Confusion Matrix for Multiple Class	27
3.5	Experimental Results and Discussion	29
3.6	Conclusion	29
4.	Conclusions & Future Work	33
4.1	Conclusions	34
4.2	Future Work	34
5.	References	35

List of Figures

1.1 Supervised Learning Model	5
1.2 Handwritten Digit Recognition System	6
1.3 Images from MNIST dataset	7
2.1 Flowchart for HOG feature extraction	11
2.2 Filter Mask for Gradient	11
2.3 Block for HOG descriptor	12
2.4 Quantization of Gradient into Orientation bins	13
2.5 Histogram of oriented gradient for 1 Block	13
2.6 Structure of HOG feature vector	14
2.7 Feature dimension vs. Block size	15
2.8 Performance of Recognition System for different Block size	16
2.9 Performance of Recognition System for different Bin size	17
2.10 Gradient and Orientation for digit 6.	18
3.1 Proximal Support Vector Machine	21
3.2 Training PSVM using One against all Method	24
3.3 Recognition Phase of handwritten digit	25
3.4 Confusion Matrix for Multiple Classes	27
3.5 Misclassification error	29
3.6 Individual Digit Recognition Rate	31

List of Tables

1.1 Dataset Distribution	8
2.1 Parameters for Feature Vector	19
3.1 Confusion Matrix for Binary Class	26
3.2 Confusion Matrix for Testing Samples	30
3.3 Design of ANN	31
3.4 Comparison between Classifiers	32

Chapter 1

Introduction

1.1 INTRODUCTION

Handwritten Digit Recognition System involves reception and interpretation of handwritten digits by a machine. Handwritten Recognition System can be divided into Offline and Online Recognition System. In case of Offline Recognition System the image of the handwritten text is scanned off- line from document by optical scanning known as optical character recognition (O.C.R.) .Where as in case of Online Recognition the text as it is written on a special digitizer or Personal Digital Assistant (PDA).

Writing styles differs in shape and orientation from person to person that makes handwriting digit recognition a challenging task. For the development of reliable handwritten digit recognition, two steps are important. The first step is extraction of discriminating feature from handwritten images and the second method is the classification of new digit images. The dimension of feature should be small. The feature to be extracted should have minimum variance within a class and maximum variance between classes. The Classifier to be used should able to classify digit with high accuracy and should take less training time during classification.

This thesis focuses on Offline Recognition System. The handwritten images are taken from MNIST (Mixed National Institute of Standards and Technology) database. The handwritten digits from 0 to 9 are trained and then tested using supervised machine learning model. Histogram of Oriented Gradient (HOG) based features are extracted from handwritten digits. Proximal Support Vector Machine (PSVM) classifier is used .The performance of classification is measured in terms four parameters derived from confusion matrix and training delay. The performance of Proximal Support Vector Machine (PSVM) Classifier is compared with Artificial Neural Network (ANN).

1.2 LITERATURE REVIEW

Handwritten digit Recognition has a wide area of research due to its vast applications. Recognition system can be divided into two major steps. First step is feature extraction from handwritten images and the second one is classification.

Researchers suggested different methodologies for feature extraction [4-7]. Fourier and wavelet based features are some of them [1-2]. The coefficients of Discrete Wavelet Transform DWT are shift variant, and the directional selectivity of subband images are poor. Complex Wavelet Transform (CWT) has been developed in order to overcome the drawbacks of DWT's but the problem is high dimension of feature vector i.e. 180 and 148 features for feature set 1 and feature set 2 respectively [3]. Histograms of Oriented Gradient Features is used for object detection like Human Detection [8-9], Pedestrian Detection [10], Large Scale Sign Detection [11] [13], Real-Time Detection and Recognition of Road Traffic Signs [12] . The popularity of HOG feature is due to invariance to local geometric and photometric transformations within local spatial or orientation bin size. In order to implement such property, it is used as a feature descriptor for handwritten digits.

For classification of features of handwritten digits, classifiers like ANN [16-18], k-nearest neighbours (k-NN) [19] and Support Vector Machine (SVM) [20-21] [28-29] are used. Out of these classifiers SVM is widely applicable. The main advantage of SVM classifier is high accuracy, but the classifier takes long training time. The computational time taken by Proximal SVM (PSVM) classifier is very less as compared to SVM [32]. Considering reduced training delay during classification, PSVM classifier is used in this work.

1.3 MOTIVATION OF THESIS

The importance of handwritten digit recognition is due to its vast practical applications .We need a highly reliable recognition system. Higher recognition rate on handwritten digits increases the recognition accuracy for handwritten data, existing in strings. Designing a handwritten digit

recognition system with 100% accuracy is practically impossible. A higher reliability system reduces financial loss due to misclassification error. For example in reading the amounts on cheques and billing [28]. Another application of handwritten digit recognition is automatic postal service. Other application involves digit recognition not necessarily handwritten. The techniques that are applicable for handwritten digits recognition can also be applied for printed digit recognition. Automatic License plate recognition is one among most widely used applications.

1.4 BRIEF INTRODUCTION TO SUPERVISED MACHINE LEARNING

Machine learning deals with the study and the construction of machine or system that can learn from data. It can also be defined as learning from experience .For example, a machine learning system can able to distinguish between spam and non-spam messages by learning on email messages. After training, it can differentiate new email messages into spam and non-spam folders. The need of machine learning is to make system automated. System automation reduces operation time and manual labour.

In supervised learning model along with input data, desirable response are also provided to the system [14]. The process is divided into two phases, first the training phase and second the testing phase. During the training phase the weights of the network are updated on the basis of training input and desirable response. Weight updating is done on the basis of specific optimization algorithm. On completion of training phase the weight of the network is considered to be optimum. This model is called as predictive model .The supervised learning model is shown in figure 1.1. A test dataset is applied to the network for validation after then the model is made generalized. Generalization refers to the ability to produce reasonable outputs for inputs not encountered during the training.

SUPERVISED LEARNING MODEL

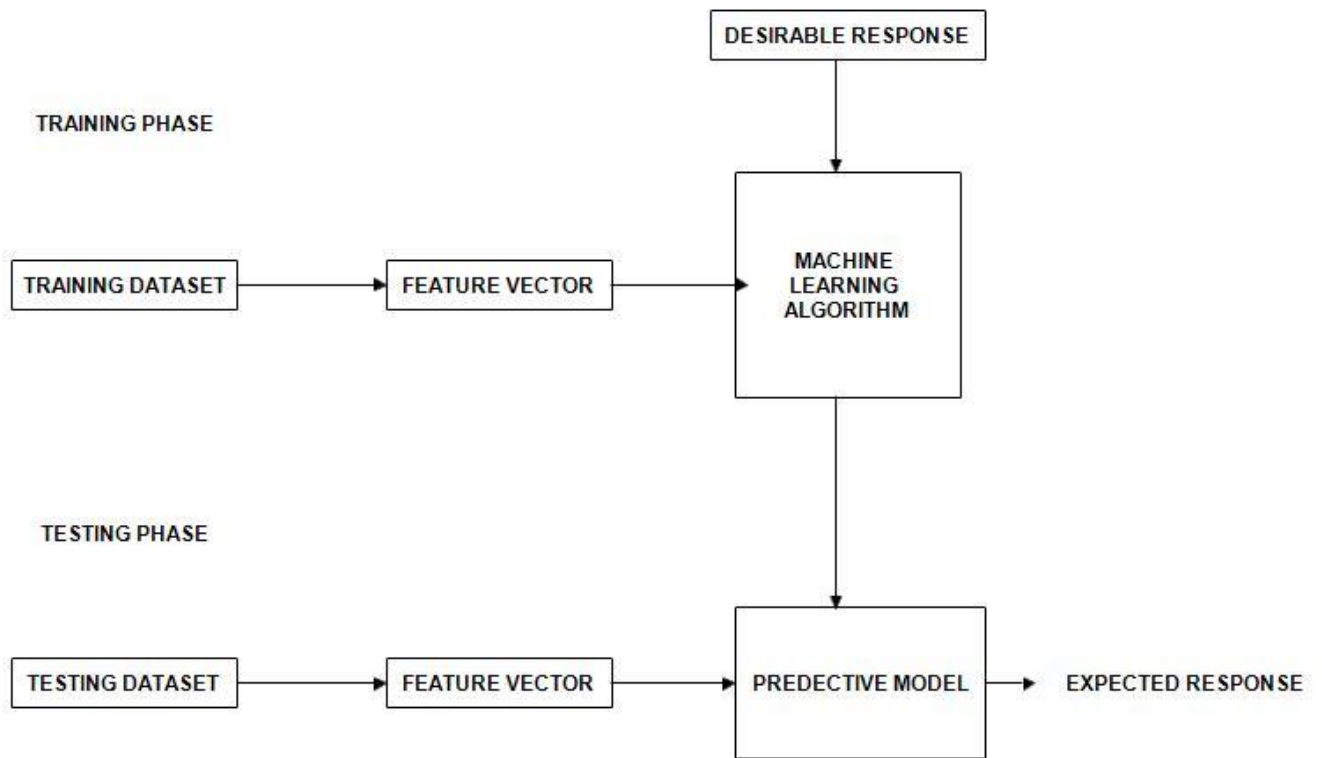


Figure 1.1: Supervised Learning Model

Supervised learning can be divided into two broad categories:

- Classification

Classification deals with the identification of a new observation's class or category, on the basis of a training dataset containing class membership information. The mathematical function which implements classification is known as classifier.

- Regression

The estimation of the relationships among variables is called as regression analysis. Forecasting and prediction are some applications of regression analysis.

1.4 SYSTEM OVERVIEW FOR HANDWRITTEN DIGIT RECOGNITION

The system overview for handwritten digit recognition system is shown in Figure no 1.2.

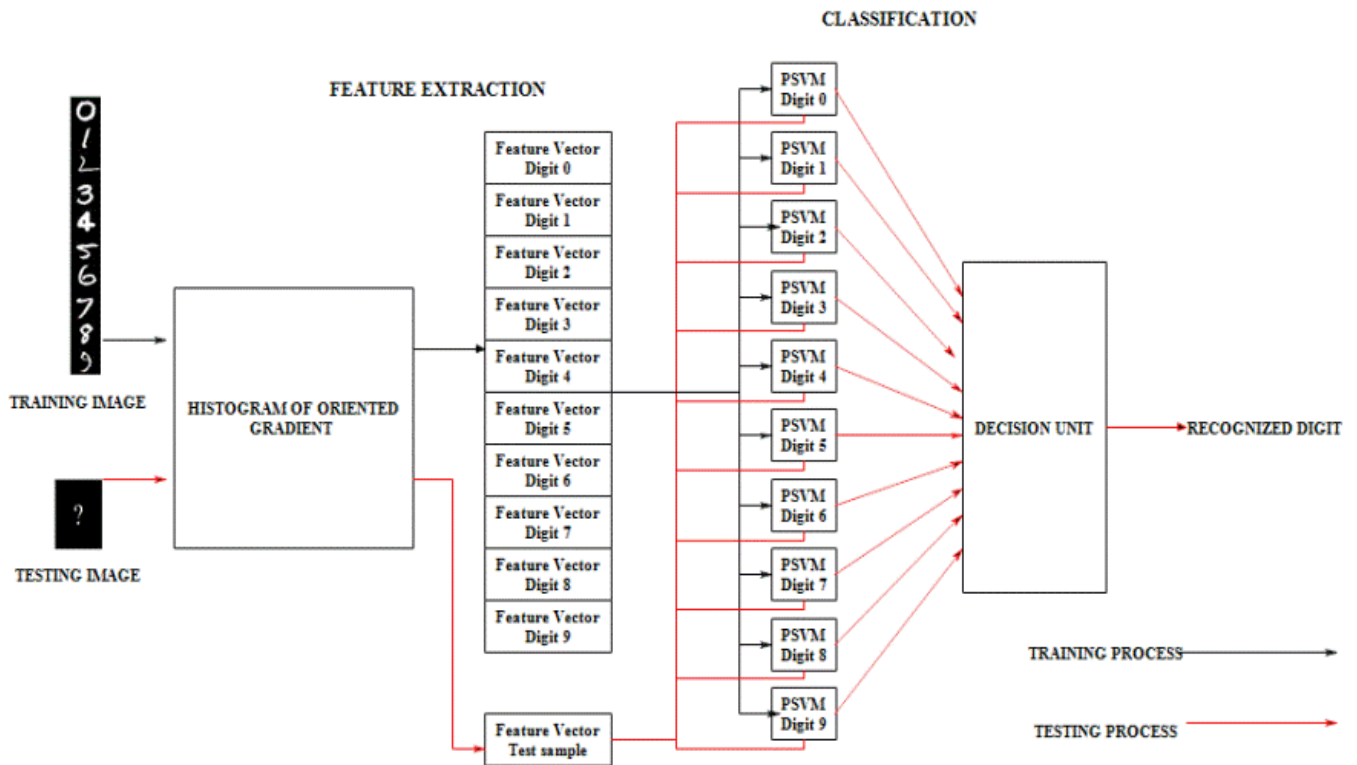


Figure 1.2: Handwritten Digit Recognition System

The learning model implemented for handwritten digit recognition is supervised. As discussed earlier, supervised machine learning has two phases- training and testing. First, from images of handwritten digits(i.e. training set images), features are extracted using HOG feature extraction .These feature vectors along with desirable response are given to PSVM networks so that the weights of the network get optimized and fixed. This completes the training phase. The entire training process is represented by black arrow as shown in figure 1.2.During testing phase; the features of unknown handwritten digits are extracted using the same feature extraction method. After then the feature vector for the test image is given to the weighted PSVM network and then the response is noted and then given to the decision unit. The decision unit makes decision and recognize the handwritten digit. In the above figure, the testing process is represented by red arrow.

1.6 DATASET DESCRIPTION

We consider the recognition system to be offline. The images of handwritten digits are taken from MNIST database and the algorithms are implemented in MATLAB codes. The training and the testing handwritten images are shown in figure 1.3



Figure 1.3: MNIST dataset

The MNIST database contains 60000 training set and 10000 testing set handwritten digits ranging from 0 to 9 [34]. Each digit is normalized and centred in a gray-level image with size 28 X28. Table no 1.1 shows the distribution of training and testing dataset for handwritten digit 0-9 that are used in this experiment. A total of 10000 samples are taken both for training and testing set respectively.

Table 1.1: Dataset Distribution

Digits	Training Set	Testing Set
0	1000	980
1	1000	1135
2	1000	1032
3	1000	1010
4	1000	982
5	1000	892
6	1000	958
7	1000	1028
8	1000	974
9	1000	1009
Total	10000	10000

1.6 THESIS OUTLINE

The thesis outline is as follows

Chapter 2: In this chapter feature extraction method for handwritten digit recognition is described. Digits are described in terms of local shape and spatial layout by means of HOG features.

Chapter 3: This chapter describes the objective function and the constraints for PSVM linear binary class classifier. The implementation of binary to 10 class classification is further mentioned. The performance of classification is measured in terms four parameters derived from confusion matrix and training delay. The performance of Proximal Support Vector Machine (PSVM) Classifier is compared with Artificial Neural Network (ANN).

Chapter 4: This chapter concludes the thesis and suggests future work.

Chapter 2

Histogram of Oriented Gradient (HOG) Based Feature Extraction

2.1 INTRODUCTION

The representation and description of images in terms of feature vector is done by means of feature extraction. The widely used feature vector for object detection in computer vision is Histogram of Oriented Gradient (HOG). Some applications of the object detection where HOG feature descriptor is used are human detection for surveillance, detection and recognition of traffic sign [8-12][31]. The aim of feature extraction is to obtain a feature vector that describes digit local shape as well as the spatial layout. The HOG feature descriptor measures the frequency of gradient orientation within local portions of a digit. Local shape and appearance of handwritten digits of an image can be represented in terms of edge directions. The local portion of the image can be obtained by dividing the image into small block. For each block histogram of gradient directions for the pixels are calculated. The combination of histograms then represents the feature vector for handwritten digits. The HOG feature vector has many advantages. As the HOG feature vector operates on localized block, it upholds invariance to geometric and photometric transformations makes small difference if they are much smaller than the local spatial or orientation bin size [8].

2.2 STEPS FOR FEATURE EXTRACTION

The feature extraction procedure from images of handwritten digit is given in figure 2.1.

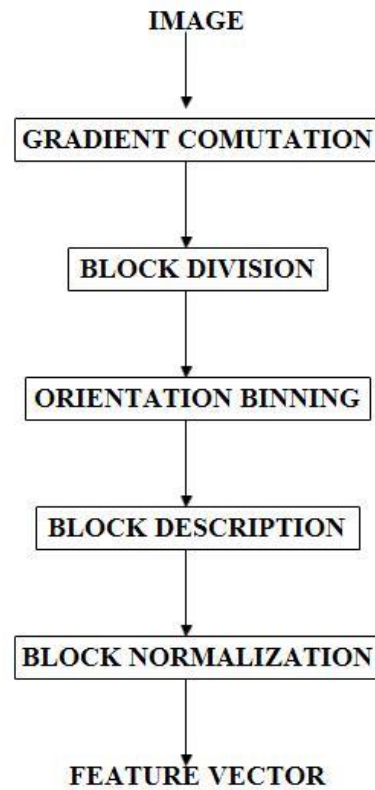


Figure 2.1: Flowchart for HOG feature extraction

2.2.1 GRADIENT COMPUTATION

This is the first step to find HOG descriptor. Gradient of an image describes edges of an object. In this step the horizontal $f_x(x,y)$ and vertical gradient $f_y(x,y)$ of image $f(x,y)$ is computed without smoothing. The filter mask along x and y direction is shown in figure 2.2.

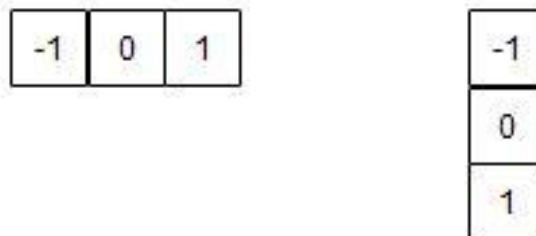


Figure 2.2: Filter Mask for Gradient

The gradient magnitude and orientation is calculated respectively as in 2.1 and 2.2

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad 2.1$$

$$\phi(x, y) = \tan^{-1} \left(\frac{f_x(x, y)}{f_y(x, y)} \right) \quad 2.2$$

2.2.2 BLOCK DIVISION

Then the image is divided into block in such way that 2 block share 50% common area of the image. Each block presents the local area within orientation of gradient directions for the pixels are computed. In figure 2.3 the rectangular portion with red and green boundaries represents 2 blocks with overlapping areas. The contribution of overlapping portion in the feature vector is more than once. In figure 2.3 the entire image is divide into a total number of 3X3=9 blocks.

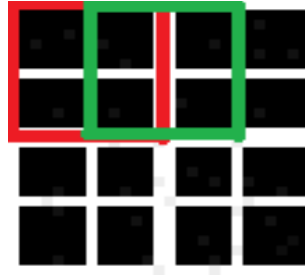


Figure 2.3: Block for HOG descriptor

2.2.3 ORIENTATION BINNING

In each block the gradient orientation is quantized into 9 orientation bin. The orientation bin are 40° equally spaced between 0° - 360° (signed gradient) [8]. The histogram/features are obtained by gradient counting into 9 orientation bins in 0° - 360° as shown in figure 2.4.

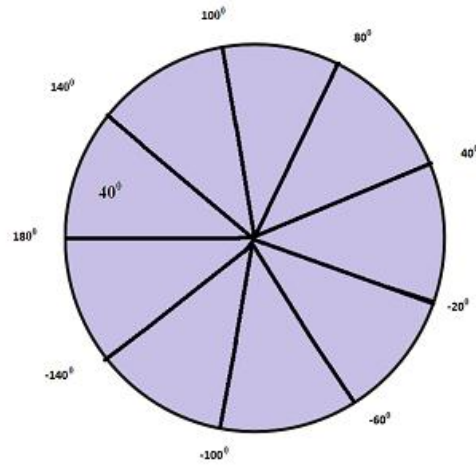


Figure 2.4: Quantization of Gradient into Orientation bins

2.2.4 BLOCK DESCRIPTION

The magnitude gradient of those pixels whose alignment lies within a particular orientation bin is summed up within a block. This is the strength of edge gradient for each orientation bin for each block. In this case each block is described in terms of histogram of 9 orientation bin as shown in figure 2.5

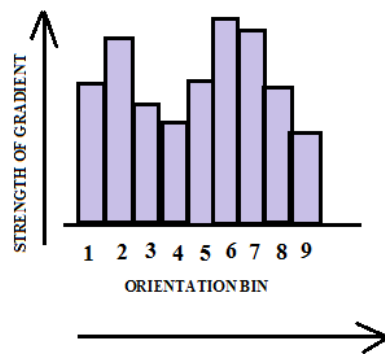


Figure 2.5: Histogram of oriented gradient for 1 Block.

2.2.5 BLOCK NORMALIZATION

The gradient strengths were locally normalized each block, in order to account for changes in illumination and contrast.

2.3 FEATURE DIMENSION CALCULATION

The dimension of feature vector is an important parameter in machine learning. Feature vector with large dimension requires large memory space for storage. For HOG feature descriptor the dimension is calculated as per 2.3

$$\text{Dimension of Feature Vector} = \text{number of Orientation Bin} \times \text{number of Blocks.} \quad 2.3$$

In this paper the number of orientation bin is 9 and the image is divided into 9 localized portions i.e. 9 blocks. So the dimension of feature vector is $9 \times 9 = 81$ as shown in figure 3.6. Hence by HOG based feature extraction method, a 2-dimensional image can be represented in terms of a 1-dimensional feature vector.

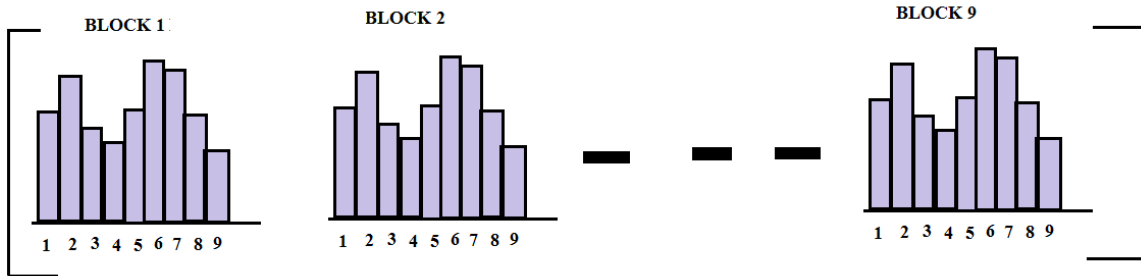


Figure 2.6: Structure of HOG feature vector

2.4 EXPERIMENTAL RESULT AND DISCUSSION

Feature extraction is an important process .While extracting feature vector for representation of handwritten images, the dimension of feature is also considered so that we do not required any additional dimension reduction technique .The dimension of feature vector depends upon block size and orientation bin size.

2.4.1 BLOCK SIZE

Consider the parameter number of overlapping block. In this work block size of 7X7 is used i.e. the size of the block is one- fourth of the image so that we can able to get 3 overlapping block with 50 % overlapping area along horizontal as well as in vertical direction. A total of $3 \times 3 = 9$ overlapping block is possible to cover entire 28X28 image.

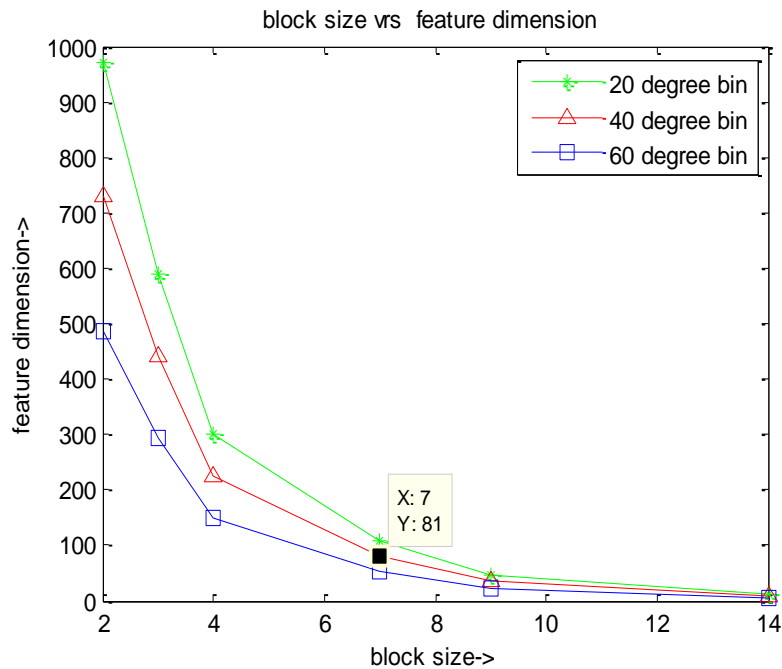


Figure 2.7: Feature dimension vs. Block size

Figure 2.7 shows the dependency of feature dimension on block size for different size of orientation bin. Block with small size requires large number of overlapping blocks and hence the dimension of feature vector is large. For large block size the dimension of feature vector is small but we cannot

take such block size as the parameters measuring the performance of the system reduces as we increase block size as shown in figure 2.8

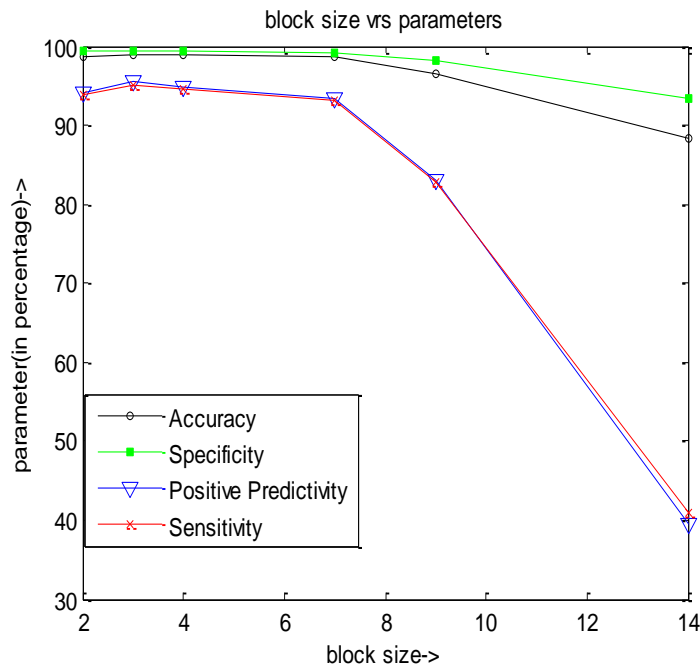


Figure 2.8: Performance of Recognition System for different Block size

In figure 2.8 shows that the parameter measuring performance of the system reduces as we increase block size. For example if block size 7 X7 the value of specificity is 99.25% which reduces to 93 % as the size of block increased to 14X14.

2.4.2. ORIENTATION BIN

Now consider the number of orientation bin. In each block the gradient orientation is to be quantized. For small orientation bin size the parameters measuring the performance of the system is high as shown in figure 2.9 ,but the no of quantization level i.e. the no of orientation bin is also high ,hence the dimension of feature vector is high.

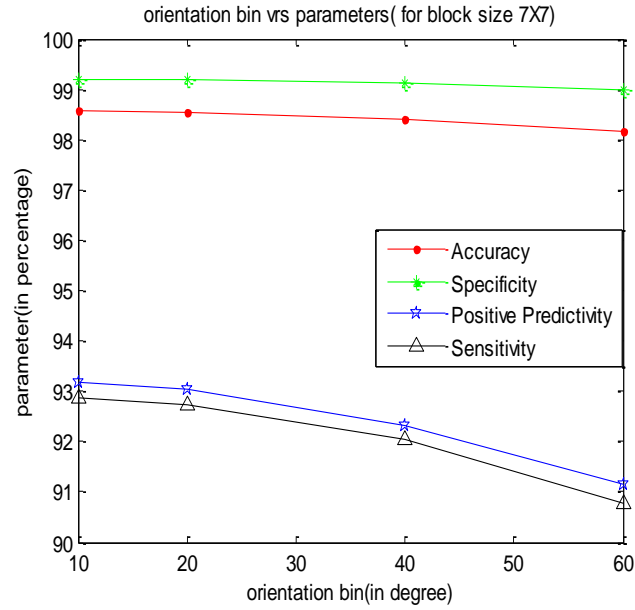


Figure 2.9: Performance of Recognition System for different Bin size

2.4.3 CALCULATION OF FEATURE DIMENSION

For orientation bin size of 10^0 the no of bins are 36.

Dimension of Feature Vector = number of Orientation Bin X number of Blocks.

For block size 7X7, the total no of overlapping blocks=9

So dimension of Feature Vector=9X36=324. The dimension of feature vector for orientation bin size of 10^0 is 324 which is large .So an extra dimension reduction technique is required.

But for 40^0 bin size, the no of bins is 9. So dimension of Feature Vector=9X9=81 which is small as well as the performance of the system is high.

2.5 CONCLUSION

Our choice of block size is 7X7 and orientation bin size 40 degree, for which the dimension of feature vector is 81. The choice of parameters are summarized in table 2.1

Table 2.1: Parameters for Feature Vector

Block Size	No. of Blocks	Orientation Bin	No. of Bins	Dimension of Features
7X7	9	40 ⁰	9	81

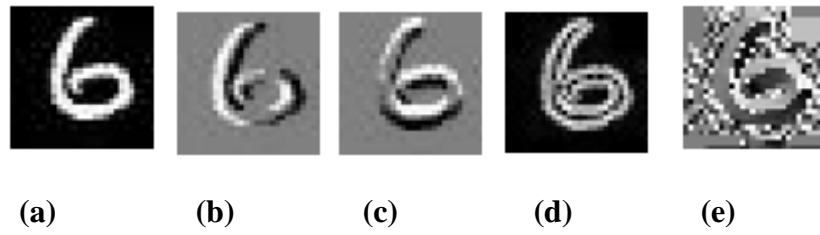


Figure 2.10: Gradient and Orientation for digit 6.

Figure 2.10(a) is the handwritten digit 6. The horizontal gradient, vertical gradient, magnitude and orientation for 6 is shown in 2.10(b), 2.10(c), 2.10(d), 2.10(e) respectively.

The final feature vector after HOG based feature extraction for 6 can be represented as

Feature Vector For Digit 6 =

[0.5000 0.1368 0.0281 0.0190 0.4580 0.6958 0.1778 0.0208 0.0644 ,
0.4121 0.2825 0.1766 0.0162 0.3224 0.5299 0.5599 0.0753 0.1224 , 0.0930
0.3136 0.3141 0.0219 0.0227 0.0127 0.8013 0.3021 0.2445 , 0.4394 0.0998
0.1336 0.3188 0.7108 0.2481 0.2133 0.1528 0.2047, 0.3025 0.2572 0.2761
0.2289 0.4537 0.3112 0.6119 0.1324 0.1381, 0.2229 0.4159 0.3210 0.0518
0.1418 0.2373 0.7149 0.1825 0.2254 , 0.1341 0.0838 0.4103 0.6165 0.5314
0.0083 0.1792 0.2314 0.2419, 0.1122 0.4453 0.5015 0.4142 0.3341 0.2177
0.4067 0.1378 0.1506, 0.2714 0.6848 0.3738 0.0584 0.1795 0.3013 0.3933
0.0077 0.1909]

1X81

Chapter 3

Multiclass- Linear Proximal Support Vector Machine (PSVM) Classification

3.1 INTRODUCTION

A Proximal Support Vector Machine (PSVM) Binary Classifier classifies features depending on the proximity to one of the two parallel plane in such a way that the distance between 2 features is maximum i.e. error of misclassification is minimum .The objective function defined for PSVM is strongly convex as a result of which the training times require for such classifier is lesser as compared of standard Support Vector Machine (SVM) Classifier[15][32].Due to this reason a Linear PSVM Classifier is preferable over Linear SVM classifier for handwritten digit recognition.

3.2 LINEAR PROXIMAL SUPPORT VECTOR MACHINE FOR BINARYCLASSES

Consider two classes A^+ and A^- . Objective is to classify m features in the n -dimensional real space R^n . The feature vector is represented by an $m \times n$ matrix A . D is $m \times m$ diagonal matrix with $+1$ s representing class A^+ and -1 s representing class A^- , along its diagonal. The quadratic programming parameter and the bias are represented by $\mu > 0$ & γ respectively. y is the slack/error variable and the column vector e of values 1 of size $m \times 1$.

The separating surface for SVM is given in equation 3.1

$$x^T w - \gamma \begin{cases} > 0, x \in A^+ \\ < 0, x \in A^- \\ = 0, x \in A^+ \text{ or } x \in A^- \end{cases} \quad 3.1$$

Let $T = \{(x_i, d_i)\}_{i=1}^m, x_i \in R^n$ be the training sample, the objective is to find the optimum values of the weight vector w_o and bias b_o such that they satisfy the constraints

$$D(Aw - ey) + y \geq e_i \text{ for } i = 1, 2, 3, \dots, m \quad 3.2$$

$$y_i \geq 0 \text{ for all } i.$$

the weight vector w , bias γ and slack variable y minimizes the cost function:

$$\phi(w, \gamma, y) = \frac{1}{2}(w^T w + \gamma^2) + \mu \frac{1}{2} \|y\|^2 \quad 3.3$$

In above equation the square norm of y is minimized. Whereas in case of SVM only the norm of slack variable is minimized. In PSVM, the square norm of y leads to strong convexity of the objective function. The inequality constraint is converted to equality as follows

$$D(Aw - e\gamma) + y = e_i \text{ for } i = 1, 2, 3, \dots, m \quad 3.4$$

$$y_i \geq 0 \text{ for all } i.$$

minimizing the cost function (objective function)

$$\phi(w, \gamma, y) = \frac{1}{2}(w^T w + \gamma^2) + \mu \frac{1}{2} \|y\|^2 \quad 3.5$$

The PSVM is shown in figure 3.1.

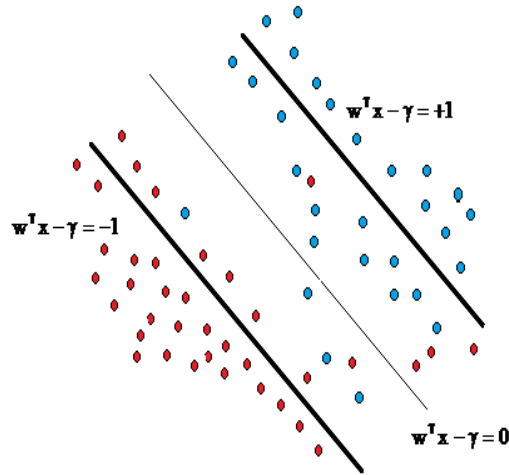


Fig 3.1: Proximal Support Vector Machine

The objective function is defined as L , where u is Lagrangian Multiplier in equation 3.6

$$L(w, \gamma, y, u) = \frac{\mu}{2} \|y\|^2 + \frac{1}{2} \left\| \begin{bmatrix} w \\ \gamma \end{bmatrix} \right\|^2 - u^T (D(Aw - e\gamma) + y - e) \quad 3.6$$

By taking derivative of L with respect to (w, γ, y, u) , the Karush-Kuhn-Tucker (KKT) necessary and sufficient optimality conditions are obtained.

$$w - A^T Du = 0$$

$$\gamma + e^T Du = 0$$

$$\mu y - u = 0$$

$$(D(Aw - e\gamma) + y - e) = 0 \quad 3.7$$

The first three equations when represented in terms of the Lagrange multiplier give the following expressions for

$$w = A^T Du$$

$$\gamma = -e^T Du \quad 3.8$$

$$y = \frac{u}{\mu}$$

Substituting the values of w , γ , and y in fourth expression of equation 3.7 and expressing u in terms of the A and D is represented as

$$u = \left(\frac{I}{\mu} + D(AA^T + ee^T)D \right)^{-1} e$$

$$u = \left(\frac{I}{\mu} + HH^T \right)^{-1} e \quad 3.9$$

where $H = D[A - e]$.

In order to simplify calculation we are using Sherman-Morrison-Woodbury matrix inversion formula, so that we get matrix of order $(n + 1) \times (n + 1)$ as follows

$$u = \mu \left(I - H \left(\frac{I}{\mu} + H^T H \right)^{-1} H^T \right) e \quad 3.10$$

ALGORITHM

In order to classify m feature vector of handwritten digit of n dimension, a matrix A of size $m \times n$ is constructed along with diagonal matrix D of labels ± 1 . The labels representing the classes of m feature vectors. Linear PSVM classifier for binary class classification is constructed as follows:

- (i) H is defined as $D[A - e]$ where e is an $m \times 1$ vector of 1s and u is computed from equation 3.10 for positive μ .
- (ii) w and γ is calculated using equation 3.8
- (iii) A new feature is then classified using PSVM classifier

3.3 IMPLEMENTATION OF MULTICLASS PSVM

The algorithm mention above is for 2 class classification. But for handwritten digit recognition system, we need a classifier that can classify handwritten digits in 10 classes i.e. from 0 to 9. One way to implement a multiclass classifier by converting a multiclass problem to many binary class problems and using many binary classifiers. Several methods are proposed for conversion of multi-class problem into several binary problems using Support Vector Machines as binary classifiers [24-27] .Such methods are listed below

- (i) One-against-all
- (ii) One-against-one

For handwritten digit classification One-against-all method is applied to binary linear PSVM classifier as the one against all significantly more accurate for digit recognition using SVM classifier [30].

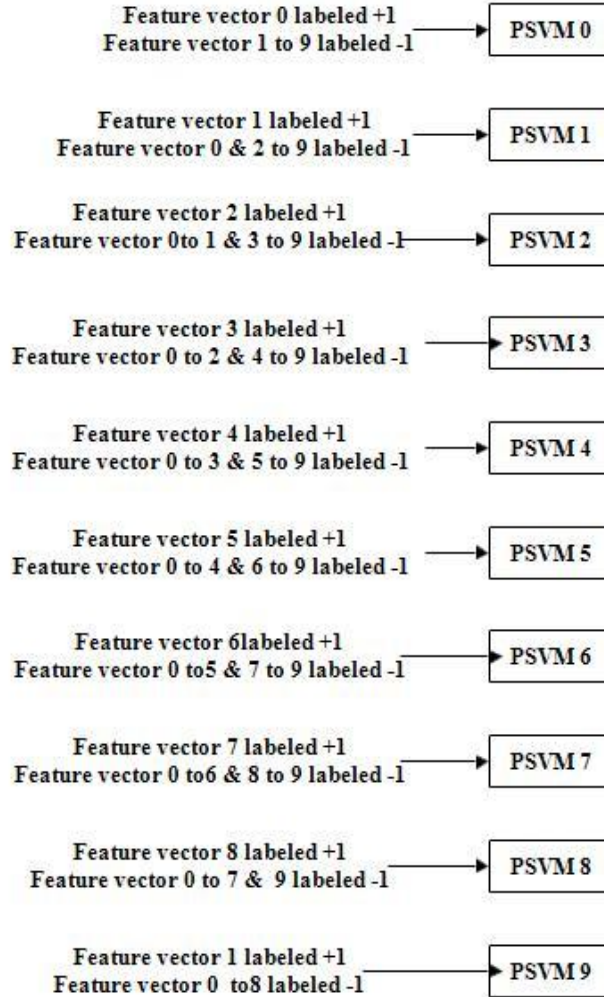


Figure 3.2: Training PSVM using One-against-all Method

3.3.1 One-against-all

In One-against-all method for the 10-class problem, 10 binary linear PSVM classifiers are constructed. The i^{th} class PSVM is trained while labeling the feature descriptor in the i^{th} class as +1 and all the rest as -1 as shown in Figure 3.2

In the recognition phase, a test feature descriptor is presented to all 10 PSVMs and is labeled according to the output which is proximal or nearest to the label +1 among the 10 classifiers in decision unit as shown in figure 3.3.

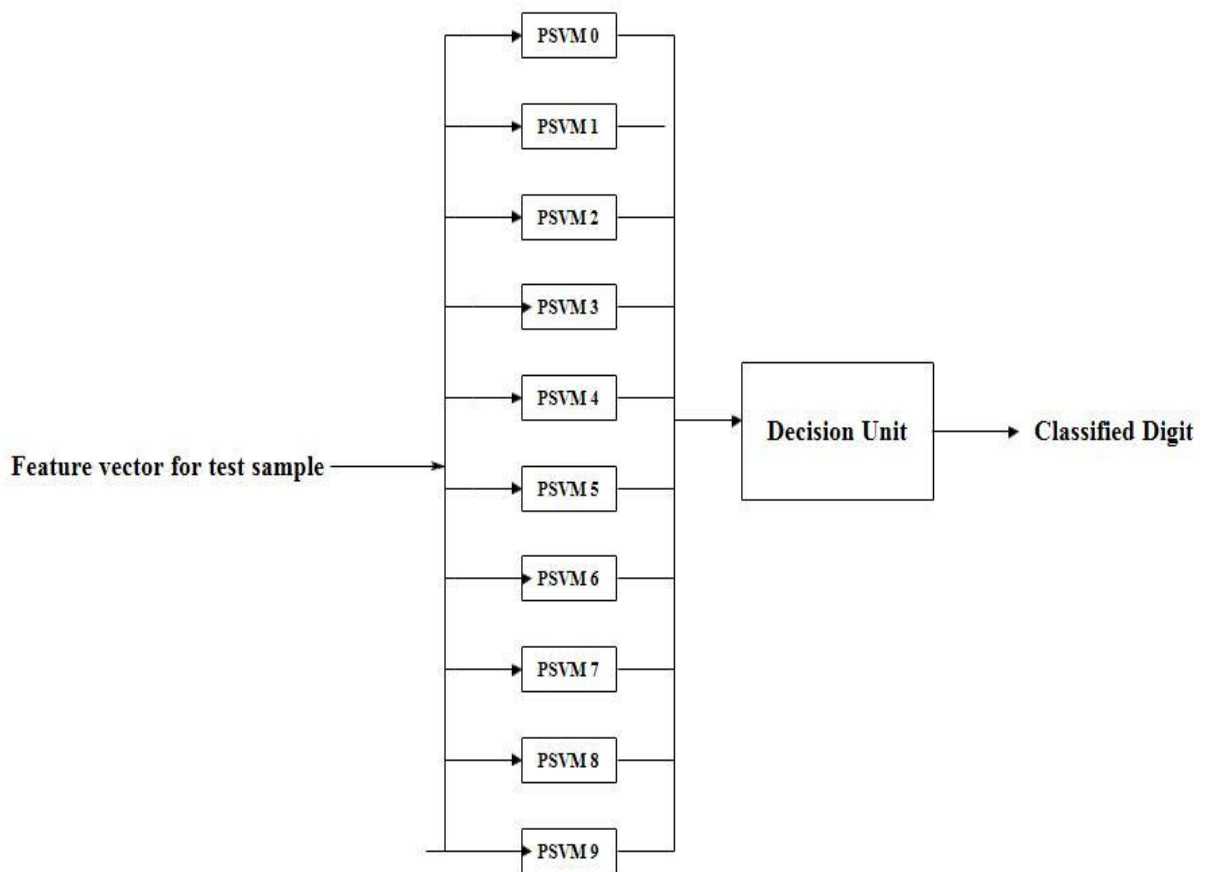


Figure 3.3: Recognition Phase of handwritten digit

3.4 CONFUSION MATRIX

In supervised machine learning, in order to show the performance of a classifier a specific table layout is drawn which is known as confusion matrix or error matrix. This matrix represents actual and predicted classifications. Each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class [31].

3.4.1 CONFUSION MATRIX FOR BINARY CLASS

Suppose we want to categorize students in 2 classes. The students securing greater than 65% mark are grouped in Class-A and the rest of the students in Class-B. The table of confusion drawn for such classification is shown below

Table 3.1: Confusion Matrix for Binary Class

CONFUSION MATRIX		Predicted	
		Class-A	Class-B
		Positive	Negative
Actual	Class-A Positive	w True positive	x False negative
	Class-B Negative	y False positive	z True Negative

The entries of the confusion matrix are listed below

w= student belonging to Class-A, correctly classified as Class-A.

x=student belonging to Class-A, but misclassified as Class-B.

y== student belonging to Class-B, but misclassified as Class-A.

z= student belonging to Class-B, correctly classified as Class-B.

The statistical measures of the performance of a binary classification are

Sensitivity: Sensitivity is the classifier's ability to identify a condition correctly.

$$\text{Sensitivity} = \frac{w}{w + x} \quad 3.1$$

Specificity: Specificity is the classifier's ability to exclude a condition correctly.

$$\text{Specificity} = \frac{z}{y + z} \quad 3.2$$

Accuracy : Accuracy is the proportion of true results (both true positives and true negatives) in the population.

$$\text{Accuracy} = \frac{w + z}{w + x + y + z} \quad 3.3$$

Positive Predictivity: For a small positive predictive value indicates that many of the positive results from classification procedure are false positives and vise versa.

$$\text{Positive Predictivity} = \frac{w}{w + y} \quad 3.4$$

3.4.2 CONFUSION MATRIX FOR MULTIPLE CLASSES

The confusion matrix drawn for multiple classes (i.e. more than 2 classes) is shown in figure 3.4.

The higher values are distributed along the diagonal of matrix.

		PREDICTED		
		A	B	C
ACTUAL	A	496	3	1
	B	10	486	4
	C	30	20	450

Figure 3.4: Confusion Matrix for Multiple Classes

In order to draw parameters from confusion table for multiple classes, the following algorithm is implemented

Given

- Confusion Matrix [CT]for multiple Classes,
- $CT(i,j)$ denotes element of confusion matrix of i^{th} row and j^{th} column, : denotes all

for $i=1$ to 10 % for each digit class

Calculate

- True Positive $TP(i) = CT(i,i)$
- False Negative $FN(i) = \text{sum}(CT(i,:)) - TP(i);$
- False Positive $FP(i) = \text{sum}(CT(:,i)) - TP(i);$
- True negative $TN(i) = \text{sum}(\text{sum}(CTt(:, :))) - (TP(i) + FN(i) + FP(i));$

Calculate Sensitivity $SEN(i)$ % equation 5.1

Specificity $SP(i)$ % equation 5.2

Accuracy $ACC(i)$ % equation 5.3

Positive Predictivity $PP(i)$ % equation 5.4

end

Calculate Mean of Sensitivity(SEN), Specificity(SP), Accuracy(ACC) and
Positive Predictivity(PP).

3.5 EXPERIMENTAL RESULT AND DISCUSSION

In this thesis first HOG based features are extracted from handwritten digits after than 10- class PSVM Classifier is used. Many handwritten digit classification system uses Support Vector Machine as a classifier .In this paper we are taking Proximal SVM as a classifier, as it requires lesser computation time than SVM for classification[32] The binary PSVM classifier is implemented using toolbox [33], which is further converted to a 10-class classifier.

Some misclassification errors are shown in figure 3.5.The number on left hand side of the equality represents correct digit and on the right hand side represents misclassified digit.

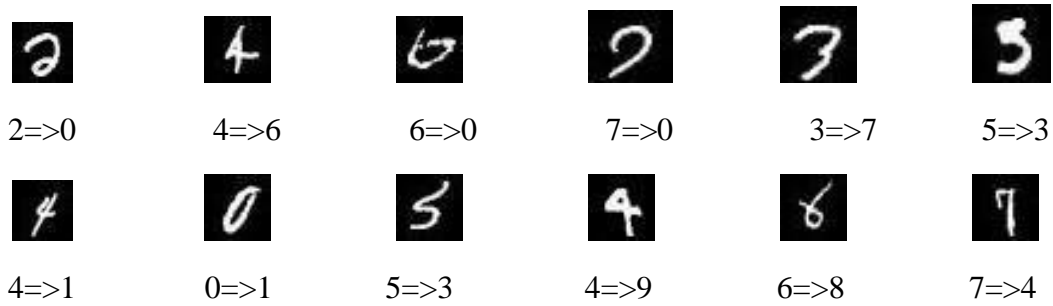


Figure 3.5: Misclassification error

Table 3.2 shows the confusion matrix for handwritten digit classification using HOG based Feature Extraction & PSVM 10 class classifier. Horizontal row represents the handwritten digits that is input or given to the recognition system and vertical column represents the classified response of the recognition system. The number belonging to 6th row and 0th column represents, 10 number of handwritten digits belonging to digit 6 but misclassified as digit 0.The number of correct classification are along the diagonal of the confusion matrix i.e. 7th row and 7th column represents the correct classification of digit 7.So total 886 out of 1028 digit are correctly classified for digit 7.

Table 3.2: Confusion Matrix for Testing Samples

Confusion Matrix		Predicted Digits									
		0	1	2	3	4	5	6	7	8	9
Actual Digits	0	971	3	2	0	2	1	1	0	0	0
	1	0	1121	8	2	0	0	2	0	1	1
	2	31	1	959	15	0	2	1	12	10	1
	3	5	0	9	951	1	15	3	14	7	5
	4	0	11	3	0	916	0	15	2	3	32
	5	2	1	3	34	0	827	7	2	12	4
	6	10	4	3	0	6	12	922	0	1	0
	7	6	13	42	7	7	1	3	886	14	49
	8	18	10	16	8	12	11	21	14	852	12
	9	6	4	4	10	21	5	1	24	12	922

The overall performance of the recognition system is

Sensitivity = 93.22%

Positive Predictivity =93.27%

Specificity =99.25%

Accuracy =98.65%

The recognition rate for individual digits are shown in figure 3.6

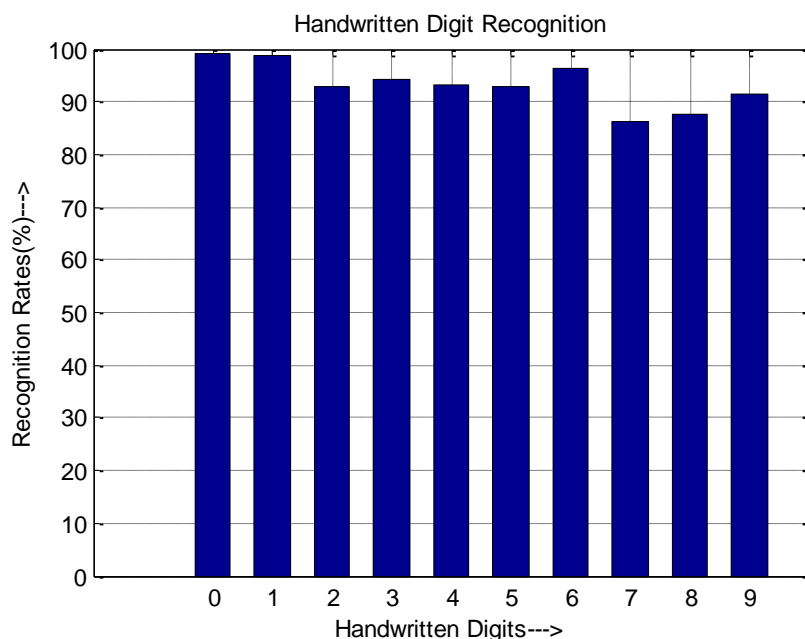


Figure 3.6: Individual Digit Recognition Rate

In the above figure, recognition rate for digit 0 is maximum and for digit 7 it is minimum. From the confusion matrix from table 3.2, it is found that most of the handwritten digits 7 are misclassified as digit 2 and digit 9. The system shows a poor recognition for digits 7 and 8 but showing a very good response for digits 0, 1, 2, 3, 4, 5, 6 and 9. The digit recognition system is also implemented using ANN classifier. The feature extraction method is same as in PSVM network. The design of 3-layer Artificial Neural Network is shown in table 3.3.

Table 3.3: Design of ANN

Training Method	Gradient Descent with Backpropagation Adaptive learning Rate
Number of Input Nodes	81
Number of Output Nodes	10
Number of Hidden Nodes	60
Activation Function	Hyperbolic tangent sigmoid
Learning Rate	0.5

Table 3.4: Comparison between Classifiers

parameters	ANN (1000 epoch)	PSVM
Sensitivity %	91.84	93.22
Positive Predictivity %	91.87	93.27
Specificity %	99.09	99.25
Accuracy %	98.37	98.65
Training Time(classifier) in second	109	0.059

The performance of Proximal Support Vector Machine (PSVM) Classifier is compared with Artificial Neural Network (ANN) in table 3.4. The overall Sensitivity, Positive Predictivity, Specificity, Accuracy for 1000 epoch ANN is 91.84, 91.87, 99.09, 98.37 respectively and by using PSVM classifier increases to 93.22, 93.27, 99.25, 98.65 respectively. The training time for ANN is 109 second which reduces to 59 milliseconds for PSVM classifier.

3.6 CONCLUSION

The Overall accuracy of PSVM classifier is found to be 98.65% with a training time of 59 milliseconds. By using PSVM classifier in handwritten digit recognition system, not only the performance is improved but also the time required to train PSVM network is reduced.

Chapter 4

Conclusions

&

Future Work

4.1 CONCLUSIONS

In this thesis HOG-PSVM handwritten digit recognition system is presented.

The images of handwritten digits are described in terms of 81 dimensions HOG feature descriptor.

The HOG feature vector holds invariance to geometric and photometric transformations that are smaller than the local region of size 7X7 or orientation bin size 40 degrees.

For recognition of HOG patterns of handwritten digits, a 10-class linear PSVM classifier is used. The overall accuracy of PSVM classifier is 98.65% with a training time of 59 milliseconds. PSVM classifier in less training time shows better performance as compared to ANN.

Along with achieving a good recognition rate of 98.65 %, the system has also maintained small dimension for feature vector (i.e. 81dimension HOG feature) without using an additional dimension reduction technique and small training time required by PSVM classifier.

4.2 FUTURE WORK

The proposed recognition system is implemented on handwritten digits taken from MNIST database.

Handwritten digit recognition system can be extended to a recognition system that can also able to recognize handwritten character and handwritten symbols. Future studies might consider on hardware implementation of recognition system.

References

- [1] Guangyi Chen and Tien D. Bui, “Invariant Fourier-Wavelet Descriptor for Pattern Recognition,” *Pattern Recognition*, Vol. 32, No. 7, pp. 1083-1088, July 1999.
- [2] Leticia M. Seijas and Enrique C. Segura, “A Wavelet-based Descriptor for Handwritten Numeral Classification,” *2012 International Conference on Frontiers in Handwriting Recognition*.
- [3] P. Zhang, T. D. Bui, C. Y. Suen, “Extraction of Hybrid Complex Wavelet Features for the Verification of Handwritten Numerals,” *Proceedings of the 9th Int’l Workshop on Frontiers in Handwriting Recognition, IEEE (IWFHR-9 2004) 2004*.
- [4] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE* 86 (11) 2278–2324, 1998.
- [5] Fabien Lauer, ChingY.Suen, Gérard Bloch, “A trainable feature extractor for handwritten digit recognition,” *Pattern Recognition*, Vol. 40, pp.1816 – 1824, 2007.
- [6] Olarik Surinta, Lambert Schomaker and Marco Wiering, “A Comparison of Feature and Pixel-based Methods for Recognizing Handwritten Bangla Digits,” *12th International Conference on Document Analysis and Recognition, 2013*.
- [7] F. Lauer, C.Y. Suen, G. Bloch, “A trainable feature extractor for handwritten digit recognition,” *Pattern Recognition*, Vol. 40 (6) ,pp. 1816–1824, June 2007.
- [8] Navneet Dalal and Bill Triggs, “Histograms of Oriented Gradients for Human Detection,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), 2005*.
- [9] Cristina Conde, Daniela Moctezuma n, Isaac Marti’n DeDiego, Enrique Cabello, “HoGG: Gabor and HoG-based human detection for surveillance in non-controlled environments,” *Neurocomputing*, Vol.100, pp.19–30, 2013.

- [10] Takuya Kobayashi, Akinori Hidaka and Takio Kurita “Selection of Histograms of Oriented Gradients Features for Pedestrian Detection,” M. Ishikawa et al. (Eds.): ICONIP 2007, Part II, LNCS 4985, pp. 598–607, 2008.Springer-Verlag Berlin Heidelberg 2008.
- [11] Gary Overett and Lars Petersson, “Large Scale Sign Detection using HOG Feature Variants,” *2011 IEEE Intelligent Vehicles Symposium (IV) Baden-Baden, Germany, June 5-9, 2011*.
- [12] Jack Greenhalgh and Majid Mirmehdi, “Real-Time Detection and Recognition of Road Traffic Signs,” *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 4, December 2012.
- [13] Fabio Boi and Lorenzo Gagliardini, “A Support Vector Machines Network for Traffic Sign Recognition,” *Proceedings of International Joint Conference on Neural Networks, San Jose, California, USA, July31-August5, 2011*.
- [14] S. Haykin, “Neural Networks—A Comprehensive Foundation,” second ed., Pearson Education, Chapter4, pp.235–240, 2006.
- [15] C. Cortes, V.N. Vapnik, “Support vector networks,” *Machine Learning* Vol.20 , pp.273–297, 1995.
- [16] Zhu Dan, Chen Xu. “The Recognition of Handwritten Digits Based on BP Neural Network and the Implementation on Android,” *Third International Conference on Intelligent System Design and Engineering Applications*, 2013.
- [17] A. Bellili, M. Gilloux, P. Gallinari, “An MLP-SVM combination architecture for offline handwritten digit recognition,” *IJDAR* Vol.5, pp. 244–252, 2003.
- [18] Zhihong Man, KevinLee , DianhuiWang, ZhenweiCao, SuiyangKhoo, “An optimal weight learning machine for handwritten digit image recognition,” *Signal Processing* ,Vol.93 ,pp. 1624–1638, 2013.

- [19] S. Impedovo, F.M. Mangini, D. Barbuizi, "A novel prototype generation technique for handwriting digit recognition," *Pattern Recognition*, Vol. 47, pp.1002–1010, 2014.
- [20] Shaileendra Kumar Shrivastava, Sanjay S. Gharde, "Support Vector Machine for Handwritten Devanagari Numeral Recognition," *International Journal of Computer Applications (0975 – 8887) Volume 7– No.11, October 2010*.
- [21] Tapan Kumar Bhowmik, Pradip Ghanty · Anandarup Roy, Swapan Kumar Parui, "SVM-based hierarchical architectures for handwritten Bangla character recognition," *IJDAR*, Vol. 12, pp. 97–108, 2009.
- [22] F. Friedrichs, C. Igel, "Evolutionary tuning of multiple SVM parameters," *Proceedings of the 12th European Symposium on Artificial Neural Networks (ESANN), Bruges, Belgium, pp. 519–524. 2004*.
- [23] O. Chapelle, V.N. Vapnik, O. Bousquet, S. Mukherjee, "Choosing multiple parameters for support vector machines," *Mach. Learn.* Vol.46(1–3), pp.131–159, 2002.
- [24] J. Weston, C. Watkins, "Multiclass support vector machines," *Technical Report CSD-TR-98-04, Royal Holloway, University of London, 1998*.
- [25] K. Crammer, Y. Singer, "On the algorithmic implementation of multiclass kernel-based vector machines," *J. Mach. Learn. Res.* 2, pp.265–292, 2001.
- [26] Y. Guermeur, A. Elisseeff, H. Paugam-Moisy, "A new multi-class SVM based on a uniform convergence result," *Proceedings of the International Joint Conference on Neural Networks, vol. 4, pp. 183–188, 2000*.
- [27] C.W. Hsu, C.J. Lin, "A comparison of methods for multi-class support vector machines," *IEEE Transactions on Neural Networks*, Vol. 13 (2) pp.415–425, 2002.
- [28] Xiao-Xiao Niu, ChingY. Suen "A novel hybrid CNN–SVM classifier for recognizing handwritten digits," *Pattern Recognition*, Vol. 45 pp.1318–1325, 2012.

- [29] Urszula, Markowska-Kaczmar, Paweł Kubacki, “Support Vector Machines in Handwritten Digits Classification,” *Proceedings of the 2005 5th International Conference on Intelligent Systems Design and Applications IEEE (ISDA'05)* 0-7695-2286-06/05 2005.
- [30] J. Manikandan, B. Venkataramani “Study and evaluation of a multi-class SVM classifier using diminishing learning technique,” *Neurocomputing*, Vol. 73, pp.1676–1685, 2010.
- [31] Oswaldo Ludwig Junior, David Delgado, Valter Goncalves, Urbano Nunes, “Trainable Classifier-Fusion Schemes: an Application to Pedestrian Detection,” *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, St. Louis, MO, USA, October 3-7, 2009*.
- [32] G. Fung and O. L. Mangasarian, “Proximal Support Vector Machine Classifiers,” *Proc. Knowledge Discovery and Data Mining, F. Provost and R. Srikant, eds., pp. 77-86, 2001*, <<ftp://ftp.cs.wisc.edu/pub/dmi/tech-reports/01-02.ps>>, Last access May 2014
- [33] G. Fung, O. L. Mangasarian, SVM toolbox home page, <http://www.cs.wisc.edu/dmi/svm/psvm>, Last access May 2014.
- [34] Y. LeCun, The MNIST database of handwritten digits, <http://yann.lecun.com/exdb/mnist/index.html>, Last access May 2014.